DRAFT PREPARED FOR SUBMISSION TO ACEGD 2017 RESEARCH CONFERENCE

PRELIMINARY AND INCOMPLETE

Trade Liberalization, Infrastructure and Firm Performance: Evidence from Ethiopia*

Matteo Fiorini† Marco Sanfilippo‡ Asha Sundaram§

August 19, 2017

Abstract

We investigate the role of infrastructure in shaping the effects of trade liberalization on firm performance. Using census data on Ethiopian manufacturing firms from 1998 through 2009 combined with information on tariff reform and improvements in road infrastructure at the town level, we show that a reduction in the tariff on inputs is associated with improvements in productivity for domestic firms. A reduction in the output tariff is associated with higher productivity for firms located in towns with better road infrastructure. Our study underlines the importance of domestic transport infrastructure in ensuring that gains from trade are spread uniformly within developing countries.

Keywords: tariffs; transport infrastructure; roads; firms; productivity; Ethiopia
JEL Classification: F14; O14; O18

1 Introduction

A large literature in international trade focuses on gains from trade liberalization for domestic firms. Trade liberalization is associated with greater firm productivity (Pavcnik, 2002; Amiti and Konings, 2007; Topalova and Khandelwal, 2011; Bigsten et al., 2016), quality improvements (Amiti and Khandelwal, 2012) and greater product scope (Goldberg et al., 2010). Similarly, there is substantial evidence that better transport infrastructure can decrease trade costs and

* We are grateful to Simona Fabrizi and members of the Galbino Project for helpful comments and conversations. We also thank IGC for funding this research. The views expressed in this paper are those of the authors. They are not meant to represent the positions or opinions of IGC. All errors are ours.

† European University Institute.

‡ University of Bari and University of Antwerp.

§ University of Auckland.

increase interregional and international trade (Donaldson, forthcoming). Better road infrastructure can spur firm activity, encourage the entry of new firms (Shiferaw et al., 2015), facilitate exports (Volpe Martincus and Blyde, 2013; Coşar and Demir, 2016) and increase employment (Volpe Martincus et al., 2017). Moreover, recent research has embedded intranational trade costs (encompassing access and quality of road infrastructure) into models of international trade and estimated the role of these intra-national barriers in shaping the pattern of comparative advantage among sub-national entities (Coşar and Fajgelbaum, 2016) as well as the intra-national distribution of the gains from falling international trade barriers (Atkin and Donaldson, 2015).

However, little has been done in the way of a formal assessment of the complementarity between road infrastructure and trade liberalization effects on firm productivity and performance. In this paper, we analyze the role played by road infrastructure in moderating the effects of a fall in import tariffs on the productivity of firms. We focus on both output and input tariffs, capturing a fall in the tariff on the final product produced by the firm and intermediate inputs used in production respectively. Reductions to output and input tariffs affect domestic product prices and hence firm productivity. While output tariff reductions lead to a lower price for the final good, spurring competition and leading to efficiency improvements, a reduction in the input tariff is associated with lower prices for intermediate inputs, allowing firms access to better quality and a wider range of inputs, boosting productivity.

We propose a framework where intermediaries operating in an imperfectly competitive market transport goods from the port to local regions in the country. Intermediaries charge a local price that includes a mark-up. In this framework, we argue that on the one hand, better road infrastructure is associated with better transmission of tariff reductions to domestic prices at the local level and hence productivity of firms in that region. Or, better road infrastructure amplifies trade liberalization effects on firm productivity. On the other hand, better road infrastructure may be associated with greater competition among intermediaries and better demand conditions in the local region. While the former amplifies trade liberalization effects on firm productivity (via its effect on the price charged by the intermediary), the latter (demand) effect operates in the opposite direction, leading to weaker transmission of tariff reductions to the local product price because the intermediary can charge a higher mark-up. Additionally, we posit that the demand effect is likely to be stronger in the case of goods used by firms as intermediate inputs, since demand in this case comes from both households consuming the good and firms using it for production. To summarize, we hypothesize that the role of road infrastructure in moderating the effects of an input and output tariff reduction depends on how each of the three effects plays out, with the demand effect exerting a stronger influence in the case of a reduction in the input tariff.

We use census data on Ethiopian manufacturing firms from 1998 through 2009.[1] The detail in our data allows us to construct measures of physical productivity at the firm level. Specifically, the availability of price data at the firm level allows us to build quantity-based productivity measures, which are not affected by the usual caveats undermining the application of revenue-

[1] The data, collected by the Ethiopian Central Statistical Agency (CSA), include detailed information on firm characteristics, location (town), labor use, capital, usage of imported inputs and exports.

based productivity in studies of the effects of trade liberalization on firm performance (see for instance De Loecker, 2011; De Loecker et al., 2016). We combine this data with two types of measures of road infrastructure quality at the town-level. The first set of measures captures accessibility, computed as the sum of the estimated travel distance that can be covered in an hour from a particular town across all roads departing from it. The second set captures the travel distance adjusted for road quality from each town to the town of Galafi, the border-town to Djibouti, which hosts the port of Djibouti handling the largest share of Ethiopia's trade. Both sets of measures account for road quality as well as the impact of road construction and rehabilitation over time.[2]

Simultaneous trade and road infrastructure reforms make Ethiopia an excellent case study for our purpose. Tariffs were reduced progressively starting in the early 1990s continuing into our sample period as part of a trade liberalization agenda initiated externally. Hence, we can argue that tariff reductions were largely exogeneous to domestic firms. Next, Ethiopia embarked on extensive improvements to road infrastructure via the Road Sector Development Programme, aimed at improving connectivity through out the country. Significant enhancements in road infrastructure were undertaken under this program, including projects to rehabiliatate and upgrade the quality of existing roads and to build new ones. The Ethiopian case thus offers an excellent setting to explore the role of road infrastructure in moderating the effects of trade liberalization.

Exploiting tariff reductions over time and improvements in road infrastructure that vary over time and across towns, our empirical strategy relates firm productivity to input and output tariffs, road infrastructure and an interaction of the two to capture the moderating effect of road infrastructure on the relationship between tariffs and firm productivity. We find that a reduction in the input tariff is assocaited with an increase in firm productivity. A ten percentage point fall in the input tariff is associated with a six percent increase in firm productivity.

We do not find evidence for a moderating role of road infrastructure in determining the relationship between input tariff reductions and firm productivity. We attribute this to conflicting effects of improved road infrastructure, arising from lower intranational transport costs, greater local intermediary competition and better local demand conditions for the intermediary. Finally, a reduction in the output tariff is associated with relatively higher firm productivity for firms with access to better road infrastructure, highlighting the moderating effect of road infrastructure in ensuring gains from trade liberalization for firms. A ten percentage point fall in the output tariff is associated with a 20 percentage point larger increase in firm productivity at the 30th percentile of road infrastructure than at the 75th percentile, demonstrating an economically significant complementarity effect. Results are robust to an instrumental variables estimation strategy, alternate measures of both productivity and road infrastructure and various cuts of the data.

Our study contributes to the literature in several ways. First, our analysis employs measures of physical firm productivity, relating to the burgeoing literature emphasizing the need to use

[2] We construct these measures from raw data gathered by the Ethiopian Road Authority (ERA), which include information on the type and quality of roads and on traffic volumes (classified according to the type and capacity of vehicles) on roads connecting Ethiopian towns.

quantity-based, rather than revenue-based measures of productivity to study the impacts of trade liberalization on firm performance. Focusing on revenue-based productivity introduces biases in the estimation of production function coefficiences and may confound the effects of trade liberalization on physical productivity and firm mark-ups. By estimating physical productivity, we are able to tackle some of these concerns. Second, we confirm the positive impact of a reduction in tariffs on intermediate inputs on physical firm productivity for firms in developing countries like Ethiopia. Third, we underscore the role for domestic transport infrastructure in ensuring gains from trade liberalization for domestic firms. Finally, we posit that in an environment where intranational trade is carried out by intermediaries with market power, the role played by road infrastructure in moderating trade liberalization effects may operate via numerous channels exerting opposing influences on local product prices and hence firm productivity.

Our paper is organized as follows. Section 2 presents our conceptual framework. Section 3 describes the empirical strategy. We describe our measures of tariffs and road infrastructure, the data and productivity estimation. We also discuss the empirical specification, identification issues and strategy. Section 4 presents results and Section 5 concludes.

2 Conceptual Framework

Our goal is to analyze the differential impact of a decrease in the output and input tariff on firm productivity in regions with differential quality of road infrastructure. We hypothesize that the impact of tariff declines on firm productivity work via two channels. First, a fall in the output tariff is associated with a fall in the product price in the local region, which results in improvements in efficiency at the firm level due to increased competition. Second, a fall in the input tariff lowers the cost of obtaining inputs (the intermediate input price), resulting in better and wider access to inputs, which boosts firm productivity. Since both channels operate via their effect on product prices, we propose a stylyzed framework to analyze the impact of tariff declines on the product price in local regions (towns in our data) in Ethiopia.

We posit a framework where intermediaries transport an imported product from the port to the final destination (Ethiopian towns, indexed by r).[3] Intermediaries obtain the product at the port for a price of $p^w + t$ where t is a specific tariff on imports of the product. They incur a transport cost to transport the product from the origin to the destination market. For simplicity, we assume that this transport cost is a per-unit cost $\tau(x_r)$ that is related to x_r, the quality of road infrastructure in region r. Better roads lead to a lower transport cost and hence, $\tau'(x_r) < 0$. Intermediaries then sell the product to households for consumption or to both households and firms (where the product is used as an intermediate input in production) at a price p_r. We conceptualize Ethiopian towns as local markets. This is consistent with anecdotal evidence from the ground, which documents how households and firms operate in local markets given high intra-national transport costs.

We assume that the number of intermediaries competing in each region r is given by n_r and it is

[3] (Atkin and Donaldson, 2015) propose a similar framework in their analysis of intra-national pass-through of world prices.

a positive function of the quality of local infrastructure x_r. Formally, $n_r'(x_r) > 0$, capturing the fact that better infrastructure induces greater competition among intermediaries. The number of intermediaries in each region is fixed (we rule out entry), which is plausible in the Ethiopian context (Atkin and Donaldson, 2015).

Consider the location specific demand function

$$Q_r = a(x_r) - bp_r \tag{1}$$

Demand in the region depends on roads via $a(x_r)$, with $a'(x_r) > 0$. This is a key assumption, unique to our framework. Our rationale behind this assumption is that with an improvement in local roads, households that were earlier unable to access this product at a nearby grocery store may now be able to access it due to a reduction in transport costs in their neighborhood. Similarly, low cost conditions generated by better access to roads might generate firm entry (or at least reduce firm exit), increasing demand for products also used as intermediate inputs in production from firms.

Denote with q_{kr} the quantity sold by intermediary k in location r. The profit function of intermediary k in location r is given by

$$\Pi_{kr}(q_1, \dots, q_{n_r}) = q_{kr} p_r\Big(\sum_{i=1}^{n_r} q_{ir}\Big) - c_{jr}(q_{jr}) \tag{2}$$

where $\sum_{i=1}^{n_r} q_{ir}$ can be denoted as q_r. Location specific marginal cost is given by

$$c_r'(q_{kr}) = p^w + t + \tau(x_r) \tag{3}$$

Necessary conditions for the Cournot-Nash equilibrium can be identified in the following system of first order conditions (FOCs):[4]

$$p_r(Q_r) - \frac{1}{b} q_{kr} - p^w - t - \tau(x_r) = 0 \qquad \forall j = 1 \dots n_r(x_r) \tag{4}$$

Summing FOCs across intermediaries, we get

$$n_r(x_r) p_r(Q_r) - \frac{1}{b} q_r - n_r(x_r)\Big(p^w + t + \tau(x_r)\Big) = 0 \tag{5}$$

Imposing market clearing $Q_r = q_r$ and using the expression for demand, (5) can be rewritten as an expression in p_r. That expression can be solved to obtain the location-specific equilibrium price

$$p_r^* = \frac{a(x_r)}{b(n_r(x_r) + 1)} + \frac{n_r(x_r)}{(n_r(x_r) + 1)}\Big(p^w + t + \tau(x_r)\Big) \tag{6}$$

Notice that the equilibrium price is equal to the marginal cost when intermediaries in the location are infinitely many. We can now derive a workable expression for $\partial\Big(\frac{\partial p}{\partial t}\frac{1}{p}\Big)/\partial x$ where, for the sake of simplicity, we ignore the r subscript since everything is now intended to be location-specific. We also write p^* as p. This expression tells us how does the proportionate change in the equilibrium

[4] Functions' differentiability and concavity of the payoff functions make FOCs also sufficient.

price associated with a change in tariffs varies with the quality of road infrastructure. In other words, it is the moderating effect of roads on the transmission of tariff changes to price changes.

$$\frac{\partial\left(\frac{\partial p}{\partial t}\frac{1}{p}\right)}{\partial x} = \frac{b}{[a + bn(p^w + t + \tau)]^2}(an' - bn^2\tau' - na') \tag{7}$$

The sign of this expression is determined by the term in the numerator in brackets.

$$\frac{\partial\left(\frac{\partial p}{\partial t}\frac{1}{p}\right)}{\partial x} < 0 \iff an' - bn^2\tau' - na' < 0 \iff \frac{n'}{n} - \frac{b}{a}n\tau' < \frac{a'}{a} \tag{8}$$

Hence, for a large enough proportional shift in demand due to better roads (the "demand" effect), $\partial\left(\frac{\partial p}{\partial t}\frac{1}{p}\right)/\partial x < 0$, suggesting that regions with better roads may see a weaker transmission of tariff declines to price declines. On the other hand, a larger competition effect among intermediaries (the "intermediary competition" effect) captured by n'/n, or a larger effect of roads on the transport cost (the "transport cost" effect) captured by $-(b/a)n\tau'$ will result in $\partial\left(\frac{\partial p}{\partial t}\frac{1}{p}\right)/\partial x > 0$ so that regions with better roads will see a stronger transmission of tariff declines to price declines. Finally, if these various effects offset each other, we may observe no moderating effect of roads on the impact of a tariff decrease on the destination price.

From our analysis above and the fact that we expect impacts on firm productivity to occur through product prices, we conclude that the moderating effect of road infrastructure on the impact of trade liberalization on productivity is ambiguous and depends on the relative strenghts of the "demand", "intermediary competition" and "transport cost" effects on product prices. If the "intermediary competition" and "transport cost" effects dominate, $\partial\left(\frac{\partial p}{\partial t}\frac{1}{p}\right)/\partial > 0$ and hence $\partial\left(\frac{\partial \omega}{\partial t}\right)/\partial x < 0$ where ω denotes firm productivity. If the "demand" effect dominates, we would expect $\partial\left(\frac{\partial p}{\partial t}\frac{1}{p}\right)/\partial x < 0$ and $\partial\left(\frac{\partial \omega}{\partial t}\right)/\partial x > 0$. We note that it is plausible to expect a stronger "demand" effect of road infrastructure in the case of products used as inputs in production if for such products, demand increases both from households and firms using the product for consumption and production respectively. We might then expect the moderating effect of roads on the impact of trade liberalization on firm productivity to be much weaker (or indeed zero) for a drop in the input tariffs relative to the output tariff. We test these hypotheses in our empirical analysis.

3 Empirical Framework

This section presents the ingredients of our empirical framework. As a first step, Section 3.1 discusses Ethiopian infrastructure and tariff reforms and related measures used in the analysis. Second, Section 3.2 describes the database of Ethiopian firms and the methodology we adopt to estimate total factor productivity. Our preferred approach accounts for both output price and input price biases in addition to the standard endogeneity concerns due to simultaneity in input choices. Third, Section 3.3 introduces the empirical specification used to analyze the role of infrastructure in moderating the effects of tariffs on firms productivity and discusses our identification strategy. Finally, Section 3.4 introduces the estimation sample and reports summary statistics.

3.1 Infrastructure and Tariffs in Ethiopia: Reforms and Related Measures

3.1.1 Infrastructure

Being a landlocked country with a poorly developed railway system (mainly connecting Addis Ababa to the port of Djibouti), road infrastructure represents the prevailing dominant transport mode for goods transported within Ethiopia (Iimi et al., 2017).[5] Recognizing this, Ethiopia has planned and implemented various sectoral infrastructure development programmes over the last 15 years. A major such programme is the Road Sector Development Programme (RSDP), which started in 1997 and was implemented in three phases, the last of which ended in 2011. A recent 13-year assessment by the Ethiopian Road Authority (ERA) reveals that the programme resulted in substantial improvements in road infrastructure (see Table 1).

Table 1: Number of firms in census years

Indicators	1997	2010
Proportion of Asphalt roads in Good Condition	17%	73%
Proportion of Gravel roads in Good Condition	25%	53%
Proportion of Rural roads in Good Condition	21%	53%
Proportion of Total Road network in Good Condition	22%	56%
Road Density/ 1000 sq. km	24.1	44.4
Road Density/ 1000 Population	0.46	0.58
Road Density/ 1000 sq. km (incl. community roads)	24	136.6
Road Density/ 1000 Population (incl. community roads)	0.49	1.83
Proportion of area more than 5 km from all weather road	79%	64.20%
Average distance to all weather road, km	21.4	11.3

Notes: Raw data sourced from RSDP 13 Years Performance and Phase IV: January 2011.

To identify the role of improvements in road infrastructure in moderating the relationship between trade liberalization and firm performance, we develop measures of improvements in the accessibility of select economic nodes (towns hosting firms in the census data) due to the rehabilitation, upgrading and construction of new roads from 1996 through 2009 (the start of the programme to the last year of our census data).

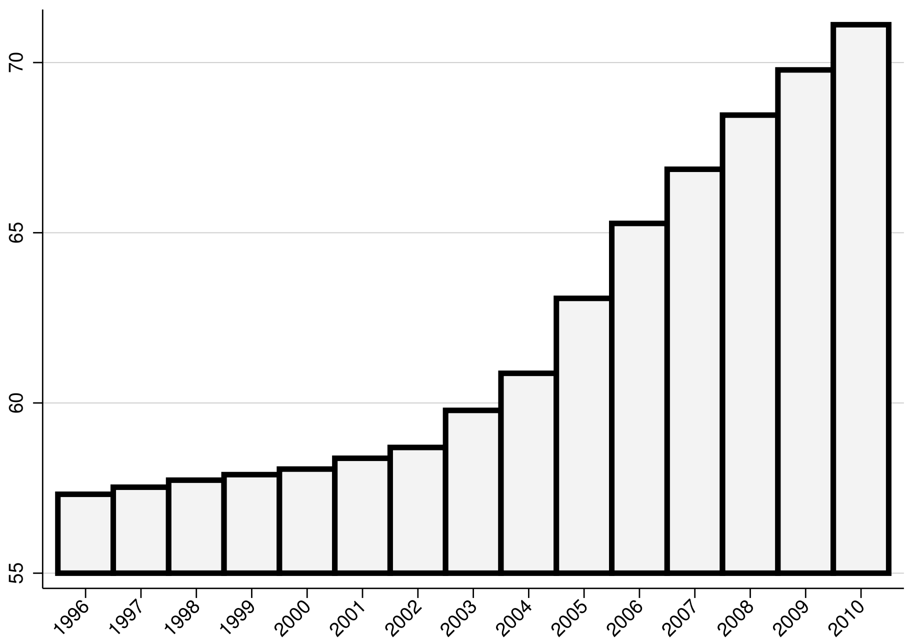
We construct measures of road infrastructure and improvements therein using data collected and updated from different sources. The main source is the 19 Year RSDP programme document, which includes a list of federal roads that were rehabilitated, upgraded or newly constructed in the past 19 years. The document includes a list of trunk roads in the rehabilitation and upgrading programme, and a list of link roads and new roads for different types of work categories such as asphalt concrete, bituminous treatment surface and gravel surface. We extract information on the year of construction and completion of each road project from physical and financial

[5] Moreover, the railway connecting Addis to Djibouti has been temporary ceased in 2007 in the section connecting Addis to Dire Dawa. The new railway connecting Addis to Djibouti has been financed by a Chinese concessional loan project and was inaugurated in early 2017.

disbursement documents. Roads are classified by their pavement type such as Asphalt Roads, Major Gravel Roads (Federal Gravel Roads), Minor Gravel Roads (Regional Rural Roads) and Earth surfaced roads. The ERA's desgin documents provide the target speed for different road classes.

The methodology implemented relies on standard analysis tools of GIS analysis. These include a service coverage analysis and a O-D (Origin-Destination) matrix analysis. Service coverage analysis tracks improvements in travel time and distance using the expected changes in travel speed due to changes in road quality. Each town included in the census data of firms represents an analysis node. We are then able to calculate the effects of road projects (new roads or road rehabilitation and upgrading) implemented on roads around the node in a particular amount of time T. We calculate two indicators. The first estimates the coverage in distance and area in one hour's time from a node to neighbouring areas using roads that branch out from this node. It is assumed that the road has an influence area or buffer zone of 12km on both sides. Figure 1 shows the overall improvement in travel distance in an hour from nodes in our firm census data over the sample period.

Figure 1: Increase in Total Travel Distance (Km) for 1hr Travel Time from Firm Node



Notes: Authors' elaboration using GIS.

The O-D matrix analysis is a tool used to test the impact of road construction or rehabilitation and upgrading on travel time to a specific destination. In the service coverage analysis described earlier, the impact of road infrastructure can only be observed if roads around the selected node are improved. The O-D matrix analysis can capture improvements in road infrastructure along road segments between the node and economic hubs. Specifically, we measure improvements in travel time to the town of Galafi (the last Ethiopian town by road to the port of Djibouti, which handles most of Ethiopia's trade).

To summarize, the measure we use as our baseline measure for the quality of road infrastructure in town r at time t is the travel distance that can be covered in an hour departing from r, summed over all roads which are accessible from r in a given year t. We denote the natural log of this measure as $\text{Infrastructure}_{rt}$. In a robustness exercise, we use travel time to Galafi from each town as an alternate measure of road infrastructure.

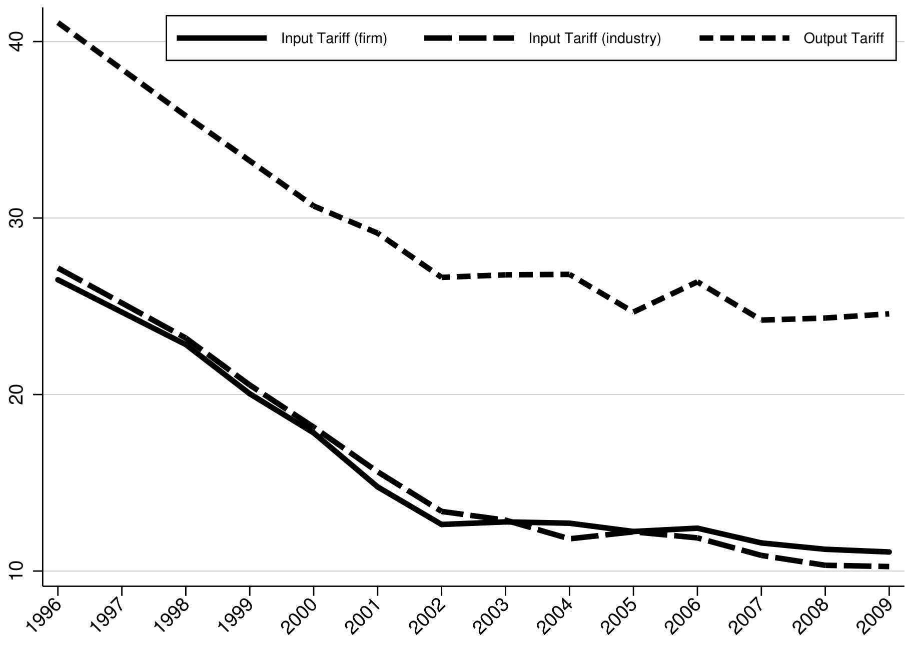
3.1.2 Tariffs

Starting in 1993, the Government of Ethiopia implemented six rounds of trade reforms, which ended in 2003 with the adoption of a six-band tariff structure with bands now ranging from 0 to 35% (more details are available in Bank, 2004).

We collect data on tariffs from the World Bank's WITS database, which uses the UN's TRAINS database as its source. Data on tariffs for Ethiopia are publicly available for the period 1995-2015, but they report some gaps in coverage, especially for the pre-2000 period. In light of this, we replace missing tariff values with values obtained by linear interpolation.

Figure 2 documents the changes in input and output tariffs for firms and sectors included in our sample from 1995 to 2009. With trade liberalization, tariffs drop consistently up to 2003 and get more stable after.

Figure 2: Changes in tariffs during the sample period



Notes: Authors' calculations from World Bank's WITS data

We construct input tariffs directly at the firm level, using information on the use of raw materials to construct weights. First, we match the code attributed by the CSA to each raw material used by the firm with a (4-digit) HS code. Second, we compute the share of each sector in each firm's total input expenditure: we denote as α_{ijt} the share of sector j input expenditure for firm i at time t.[6] Third, we use these shares as firm-specific coefficients to weight output tariffs using the standard approach.[7]

$$\text{Input-tariff}_{it} = \sum_j \alpha_{ijt} \text{Ouput-tariff}_{jt} \tag{9}$$

Input tariffs at the firm-time level and Output tariffs at the sector-time level are the trade policy variables used in the empirical analysis.

3.2 Firm-level Data and the Estimation of Total Factor Productivity

3.2.1 Firm level Data

We use firm level data from the annual census of Large and Medium Manufacturing firms, published by the Central Statistica Agency (CSA) of Ethiopia. Data cover all firms that employ at least 10 workers and that use electricity in their production process. All firms need to comply with CSA requirements, and the census is therefore representative of more structured and formal firms in the country.[8] The dataset includes detailed information on basic firm characteristics, including production, employment, capital and inputs, which are needed to estimate production functions. Firms belong to the manufacturing sector, and their industry is defined according to their 4-digit level of the ISIC Rev. 3 classification.

A key feature of the dataset is that it includes detailed information on (up to 12) specific products produced by each firm. Products are recorded following a classification made by the CSA, and for each product, information available includes the value and the quantity produced, both for the domestic and the export market. As anticipated in the discussion of tariff, our data allows us to identify raw materials used at the level of the firm and their share in total firm's expenditures. This information have been used for the construction of the input tariff variable.

Finally, and importantly for our focus on infrastructure, for each firm we have information on its region, woreda (district) up to the level of the town. While firms are located in about 90 towns in the country, and their growth is geographically diversifying over time (Mukim, 2016), we register a strong concentration in the capital, Addis Ababa, which hosts 45.7% of the firms and 52% of the total observations, respectively.

We use an unbalanced panel of 3,551 firms covering the period 1998-2009, totalling 12,672 observations. Table 2 reports the number of firms for each year of the sample, showing strong dynamism of the private sector,[9] which is consistent with the overall pattern of economic growth

[6] Importantly, we have constructed the cost shares on the basis of total input purchases, i.e. including both domestic and imported inputs, to avoid endogeneity bias (see discussion in Amiti and Konings, 2007).

[7] Note that due to the high turnover of firms, we do not set weights for the first year as usually done to reduce potential endogeneity bias from firms adapting their input mix to tariff changes.

[8] In 2005, a representative survey of firms was conducted instead of a census. This does not represent a huge bias for our analysis, since we do not focus explicitly on entry and exit rates (except when adjusting our TFP estimates for attrition), or on generating aggregate figures. Yet, we make an adjustment for those firms that are in the data in both 2004 and 2006, but not in 2005, filling in information for all the variables as the simple average of the closest years. Results remain robust when dropping 2005 from our data.

[9] The sector has experienced rapid growth, with an annual average of 10% over the period considered.

experienced by the country during the last decade (Moller, 2015).

Table 2: Number of firms in census years

year	firms	share
1998	701	5.53
1999	712	5.62
2000	704	5.56
2001	732	5.78
2002	866	6.83
2003	923	7.28
2004	980	7.73
2005	978	7.72
2006	1,131	8.93
2007	1,301	10.27
2008	1,696	13.38
2009	1,948	15.37
Total	12,672	100

Notes: Authors' elaboration on Ethiopian Census Data.

While presenting a multi-product structure, the database has a limitation which undermines the application of a multi-product empirical framework in the analysis. In particular, the product code which is necessary to identify the observation at the product-time level is missing in almost 59% of the product-time level observations in the whole sample. Other products which are not perfectly identifiable across time are those that have the same product code and the same unit of measure within a firm-year pair. These observations account for another 20% in the whole sample. On average across firms and years, non-identifiable products account for 63% of total firm sales. This would make a product-level analysis significantly non-representative.[10] For this reason we choose to conduct the TFP estimation and the subsequent econometric analysis at the level of the firm, making the necessary assumptions to exploit all the product-level information contained in the dataset. These will be discussed in the next section.

3.2.2 TFP Estimation

To construct our dependent variable, we follow the existing literature and use a measure of firm performance based on estimated Total Factor Productivity (TFP). Due to the lack of information on firm specific prices, it has been argued that a large body of research on the nexus between trade liberalization and firm performance has been unable to capture improvements in physical efficiency, but mostly captures gains in profitability. In our study, we exploit information on values and physical quantities at the product level to construct a measure of physical productivity (Eslava et al., 2004; Smeets and Warzynski, 2013).

[10] A similar issue of non-identifiability applies to raw materials which are identified with a code similar to that used for products. About 56% of all raw-material-time level observations have a missing raw material code and therefore are non identifiable across time. Moreover 12% of the observations have the same (non-missing) raw material code and the same unit measure within a firm-year pair. Fortunately the problem here is less relevant from an economic point of view as non-identifiable raw materials account only for about 30% of all raw material expenditures on average across firms and years. This is why the input tariff variable at the firm-level built on these data remains our preferred measure of input tariffs.

We start from a basic production function linking the output produced by firm i to the costs of inputs adopted in the production process:

$$y_{ijrt} = \beta_1 \ k_{ijrt} + \beta_2 \ l_{ijrt} + \beta_3 \ m_{ijrt} + \omega_{ijrt} + \epsilon_{ijrt} \tag{10}$$

where y_{ijrt} denotes the output of firm i producing in sector j, located in town r at time t. k_{ijrt} denotes capital, l_{ijrt} labour and m_{ijrt} the costs of materials, respectively. The random component ω_{ijrt} is the unobservable productivity or technical efficiency and ϵ_{ijrt} is an idiosyncratic output shock distributed as white noise.

Standard approaches adopt industry price deflators, when available, to adjust both output and inputs for price variation common to all firms in a given industry j. This introduces a so-called output price bias, resulting in a downward bias of the input coefficients, which is due to the likely correlation between the firm-specific variation in output prices and expenditures on inputs (De Loecker and Goldberg, 2014; De Loecker, 2011). Similarly, lack of information on input price variation can introduce a downward bias in the estimated coefficients, given that higher input prices will raise input expenditure while not increasing physical output (De Loecker et al., 2016).

Our data allow us to eliminate the output price bias given that we can calculate prices at the product level, since firms report information on quantity and values of products they produce (see Section 3.2.1 above).[11] We aggregate product prices at the firm level calculating a firm-level price index P_{it}, using the approach suggested by Eslava et al. (2004) and Smeets and Warzynski (2013). The steps followed to calculate P_{it} are described in the Appendix A.

While deflating output with firm specific prices eliminates the output price bias and allows us to compute physical TFP more precisely, to address the input price bias we follow a simplified version of the approach developed by De Loecker et al. (2016). The assumption here is that the source of input price variation at the firm level can be captured by the quality of inputs adopted in the production process. Another assumption is that output quality is complementary to input quality, and therefore, the quality of inputs is a function of the quality of output. With this assumption, the input price bias can be accounted for by including the output price index in the control function to account for unobserved input price variation.

We estimate production functions at the sector level (aggregating sectors at the 2 digit of the ISIC classification, and combining sectors sharing similar technologies when sample sizes are too small). Since OLS coefficients will be biased in equation (10) due to simultaneity and selection biases, we apply the approach by Levinsohn and Petrin (2003) (LP) that uses raw material inputs as a proxy for unobservable productivity shocks to correct for the simultaneity bias. We also address potential collinearity in the first stage due to simultaneity bias of the labour coefficient by adopting the correction suggested by Ackerberg et al. (2015). Finally, we adjust our estimates for attrition in the second stage of our productivity estimation. Physical output is the total production at the level of the firm deflated using P_{it} described in the Appendix A. We use the

[11] This comes with measurement issues, given that prices are measured by unit values. In addition, we make assumptions concerning product homogeneity or the way products are aggregated across firms (see De Loecker and Goldberg, 2014, for more discussion).

book value of fixed assets at the beginning of the year to estimate the capital coefficient, the total number of permanent employees for labour and the costs of raw materials for intermediate inputs.

3.3 Econometric Specification and Identification Strategy

The basic empirical strategy used in this paper consists of a standard interaction model, where the main regressor of interest is the product of the policy treatment (output or input tariff) and a moderator variable (quality of insfrastructure). The specification varies depending on the policy instrument considered, whether it is input or the output tariff. The two respective empirical models are given by

$$\log \mathrm{TFP}_{ijrt} = \beta \text{ Input-tariff}_{ijrt} + \gamma \text{ Input-tariff}_{ijrt} \times \text{Infrastructure}_{rt} + \boldsymbol{\delta}' \mathbf{z}_{ijrt} + \mu_i + \nu_{rt} + \varepsilon_{ijrt} \quad (11)$$

$$\log \mathrm{TFP}_{ijrt} = \beta \text{ Output-tariff}_{jt} + \gamma \text{ Output-tariff}_{jt} \times \text{Infrastructure}_{rt} + \boldsymbol{\delta}' \mathbf{z}_{ijrt} + \mu_i + \nu_{rt} + \varepsilon_{ijrt} \quad (12)$$

The dependent variable in both equations is the natural logarithm of TFP estimated for firm i active in sector j, town r at time t. Input tariffs in equation (11) vary at the firm level. Instead, output tariffs in equation (12) are specific to the sector j and do not vary across firms within the sector. The second regressor in both specifications consists of the interaction between the respective policy treatment and our measure of the quality of infrastructure. The latter varies at the town level and over time. Both models feature a vector of firm-specific characteristics varying over time ($\mathbf{z}_{ijrt}$): this includes a control for the firm's age (age_{ijrt}), a dummy for exporter status ($\text{Exporter dummy}_{ijrt}$) and one for foreign ownership ($\text{Foreign ownership dummy}_{ijrt}$). All specifications also contain firm fixed effects (μ_i), town-time fixed effects (ν_{rt}) and the idiosyncratic error term (ε_{ijrt}).

Consistent with the large literature on the productivity effects of tariff liberalization, lower tariffs are expected to have a positive impact on TFP at the average quality of infrastructure. This would be reflected in a negative sign for the coefficient β when the moderator variable $\text{Infrastructure}_{rt}$ is demeaned. By construction, the proposed specifications allow the productivity effect of the respective policy instrument to vary linearly with the quality of infrastructure. The role of infrastructure in shaping the effect of tariff liberalization is identified by the coefficient γ. As discussed in Section 2, the different theoretical channels determining the moderating role of infrastructure exert opposing influences. As a consequence, the sign of γ is ultimately an empirical matter. To reiterate, if the "transport cost" and "interemediary competition" effects of road infrastructure dominate the "demand" effect, the sign of γ will be negative. If the "demand" effect offsets the two other effects or if it dominates, we may observe a zero or positive sign for γ.

Identification in this empirical setting requires the policy treatments to be as good as randomly assigned in each equation. The included battery of fixed effects accounts for any confounding heterogeneity originating from firm-specific as well as town-time-specific shocks/characteristics. In sections below, we address further concerns pertaining to our empirical strategy.

3.3.1 Endogeneity of Infrastructure

Literature has long debated the potential endogeneity of infrastructure investments, claiming that new infrastructure is often placed where it will have the biggest economic impact (Coşar and Demir, 2016; Duflo and Pande, 2007). In the specific context of our analysis, this would be the case if the decision on where to place new roads or to improve existing ones is somehow related to the presence of more productive firms in given towns. While the economic potential of the area is, along with other criteria, among the priorities listed by the Ethiopian Road Authority (ERA) for allocating new infrastructure investments, there is little evidence on the effective criteria used to determine investments in practice (Shiferaw et al., 2015). The work by Shiferaw et al. (2015) based on our same data claims that the risk of endogenous road placement decisions should be marginal given that plans of road constructions are taken on a 5-year basis and this can hardly affect annual changes in firm performance. In addition, given the small weight of the manufacturing sector in the economy, it is hard to speculate that its current performance could affect long term investment decisions in the road sector.

We do not expect potential endogeneity of road placement to present a significant concern in our context. Given that we control for time-varying town fixed effects in our baseline regressions, we rule out potential simultaneity bias between firm productivity and road construction due to their possible correlation with economic and political conditions in towns. Yet, for robustness, we account for the potential endogeneity of road infrastructure by employing an instrumental variable (IV) approach. Following recent studies (Duflo and Pande, 2007; Iimi et al., 2017; Wang et al., 2016), we use the geophysical condition of the terrain as a plausible exogenous proxy for the higher costs of building road infrastructure. We use the average slope of the terrain in the district weighted by the distance of the town to the capital, Addis as an instrument for road infrastructure on the basis of the assumption that investing in new road infrastructure in areas more isolated from Addis (the capital and geographical center) is more costly. We use an interactions of this instrument with tariffs as instruments for our two interaction terms of interest. Given that we use the instrument interacted with the tariff variables that vary by year, we do not worry about the time invariant nature of the proposed instrument.

3.3.2 Endogeneity of Tariffs

A standard argument in the literature has to do with the potential endogeneity of trade policy. Political economy mechanisms (Grossman and Helpman, 1994), including the targeting of more (or less) productive industries for protection or lobbying by firms and industries might influence both the timing and the size of trade protection, introducing a bias in our estimates. In the case of Ethiopia, we are confident about the exogeneity of trade policy since, as also argued by Bigsten et al. (2016), trade reform was largely shaped by International Institutions under liberalization programmes in the early '90s. Yet, since we cannot completely rule out endogeneity of tariffs on the basis of this argument, we try to address this potential concern in three ways.

First, as done by Topalova and Khandelwal (2011), Ahsan (2013) for India and by Bas (2012) for Argentina, we aggregate our firm data at the industry level to test for the political protection argument. Specifically, we construct aggregates of production, employment, export, capital intensity and firms agglomeration for each 4-digit industry and test the correlation among pre-sample levels (1996) of these variables and changes in the input and output tariffs between 1996 and 2003.[12] Results of these regressions show that there is hardly any correlation between changes in tariffs and pre-sample industry characteristics, bolstering our argument that tariff reform in Ethiopia was largely exogeneous to firm outcomes.

Second, following Topalova and Khandelwal (2011), we check whether tariff adjustments were made in response to productivity levels. To do this we regress input tariffs (calculated at the firm level) at time $t+1$ on firm productivity at t, controlling for firm and year fixed effects. We do the same for output tariffs, regressing their level at $t+1$ on an indicator of average industry productivity, weighted by the share of firm total output and controlling for industry and year fixed effects. We repeat the same exercise using levels of productivity at $t-5$. Results of these exercises show that changes in tariffs were not correlated to previous levels of productivity of firms and industries, thus implying that policymakers did not adjust trade policy in response to observed productivity levels of local firms and industries.

Third, to check the robustness of our baseline empirical assessment of the effects of tariff liberalization at the mean value of infrastructure quality, we employ an instrumental variable (IV) approach. This IV exercise uses the pre-sample level of tariffs in 1995 (note that this does not correspond to pre-reform year, which is instead 1993, due to tariff data availability) as an instrument for long term changes in tariffs as in Amiti and Konings (2007) and Ahsan (2013). Pre-sample tariff levels should be uncorrelated with the current error term, but are certainly linked to changes in tariffs over time (given that tariffs were rationalized as part of the liberalization programme, industries with higher tariffs in 1995 should have experienced more rapid decreases in tariffs).

We discuss results from the two IV exercises after the discussion of our baseline estimates. Overall, we find that our baseline results hold qualitatively.

3.4 Estimation Sample

Assembling data from the different sources we obtain the estimation sample which consists of an unbalanced panel covering up to 1532 firms in the period 1998-2009, yielding a total of 7463 observations. Summary statistics for the variables used to obtain the baseline results are reported in Table 3.

[12] To do this exercise we use input tariffs computed at the industry, rather that at the firm level. We use the change until 2003, since this is the year of the latest trade reform. Results do not change if we replicate the same exercise using the change in tariffs from 1996 to 2009.

Table 3: Summary statistics

Variable	mean	median	sd	min	max
log TFP_{ijrt}	2.708	2.451	1.819	-7.583	12.981
$\text{Input-tariff}_{ijrt}$	13.940	10	9.934	0	66.667
$\text{Output-tariff}_{jt}$	27.182	29.460	9.764	5	65.347
$\text{Infrastructure}_{rt}$	7.601	8.330	1.088	4.361	8.423
$\log(\text{age}_{ijrt} + 1)$	2.424	2.398	0.893	0	4.736
Exporter dummy_{ijrt}	0.048	0	0.215	0	1
Foreign ownership dummy_{ijrt}	0.039	0	0.194	0	1

Notes: The table reports summary statistics for the main variables used in the analysis.

4 Results

4.1 Baseline Results

Section 2 outlines our hypothesis that the role of road infrastructure in moderating the relationship between tariff reductions and firm productivity that operates through product prices is ambiguous and depends on the relative strenghts of the “trasnport cost” and “intermediary competition” effects on the one hand and the “demand” effect on the other. If the first two effects dominate, we expect coefficients on the interaction terms between tariffs and road infrastructure to be negative, suggesting that a reduction in tariffs is associated with relatively higher productivity for firms in towns with better road infrastructure. If the latter effect dominates, we expect coefficients on the interaction terms between tariffs and road infrastructure to be positive. We also expect the likelihood of the "demand" effect dominating to be higher in the case of the input tariff relative to the output tariff, since demand for inputs comes from households and firms.

We report baseline results in Table 4. Columns (1) through (4) present results for versions of equations (11) and (12). The first two columns exclude interactions of tariffs and road infrastructure, which are then included in Columns (3) and (4). Column (5), our main specification, includes both input and output tariffs and their interactions with road infrastructure in a single regression. Hence, in Column (5), we observe input (output) tariff effects conditional on the output (input) tariff.

From Columns (1) and (2), we find that reduction in the input and output tariffs are associated with an increase in firm productivity. However, the coefficient is statistically significant only in case of the input tariff. A ten percentage point fall in the input tariff is associated with six percent increase in firm productivity. Consistent with the literature, this result lends support to the idea that better access to intermediate inputs through lower prices is associated with productivity improvements for firms. From Columns (3) and (4), negative coefficients on the interaction terms suggest that a reduction in input and output tariffs is associated with higher productivity for firms in towns with better road infrastructure. Our results point to a key role for road infrastructure in moderating the effects of trade liberalization on firm productivity. We note that though the negative coefficient on the output tariff in Column (2) is not statistically significant, results in Column (4) suggest that the impact of an output tariff reduction on firm productivity is much stronger with better road infrastructure. This emphasizes the role for road infrastructure in ensuring gains from trade liberalization for firms.

In Column (5), we find that while the signs of our coefficients of interest remain consistent with those in previous columns, it is only the interaction between the output tariff and road infrastructure that is statistically significant. This is consistent with the idea we propose in our conceptual framework. If the "demand" effect on the product price of a fall in the input tariff dominates or offsets the "transport cost" and "intermediary competition" effects, we would not observe a moderating effect of road infrastructure in the case of the input tariff.

Figure 3 presents the marginal effect of the output tariff variable for varying levels of the road infrastructure variable $\text{Infrastructure}_{rt}$. For values of $\text{Infrastructure}_{rt}$ bigger than 7.06, the marginal effect of the output tariff is negative, suggesting a positive productivity effect of output tariff liberalization. For values of $\text{Infrastructure}_{rt}$ bigger than 7.95 (roughly corresponding to the 33rd percentile), point estimates are negative and statistically different from zero. A ten percentage point fall in the output tariff is associated with a 20 percentage point larger increase in firm productivity at the 75th percentile of road infrastructure than at the 30th percentile, demonstrating an economically significant complementarity effect.

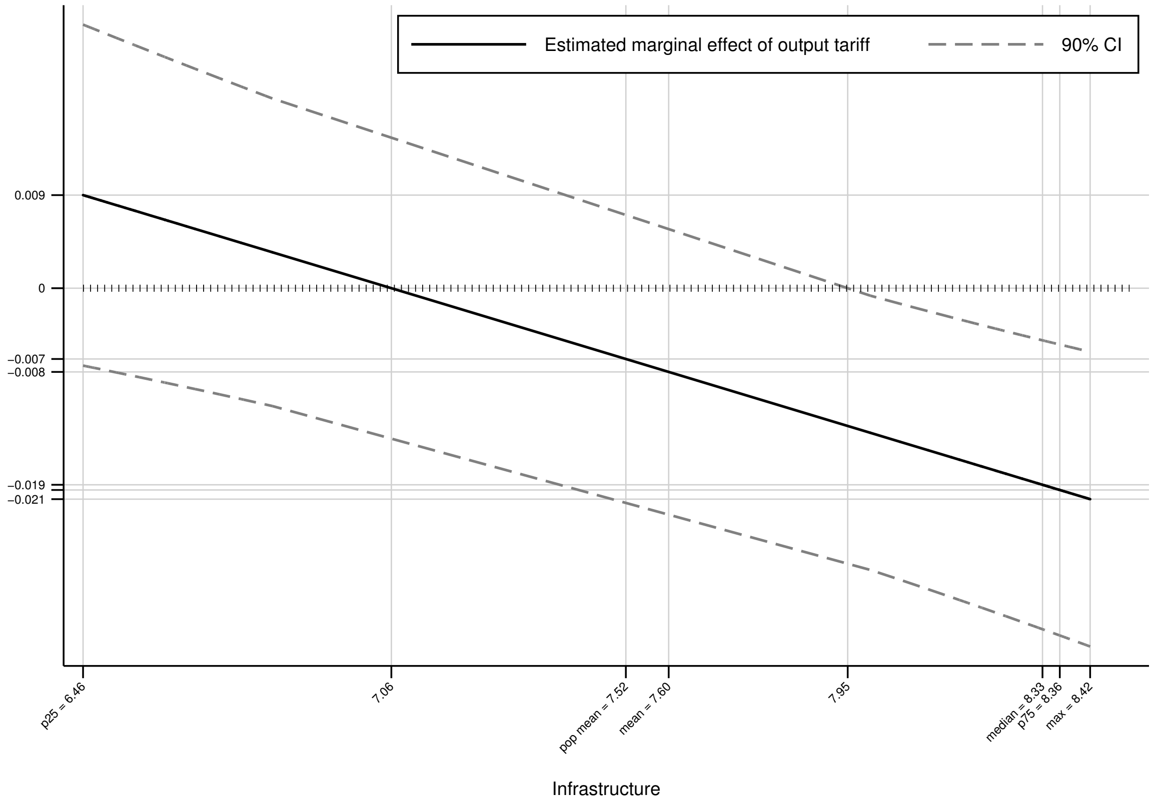
To summarize, our baseline results indicate that a reduction in the input tariff is assocaited with an increase in firm productivity. We do not find evidence for a moderating role for road infrastructure in determining trade liberalization effects on firm productivity in the case of a reduction in the input tariff. We attribute this to conflicting "demand", "transport cost" and "intermediary competition" effects of a reduction in input tariffs on product prices and therefore firm productivity. Finally, a reduction in the output tariff is associated with relatively higher firm productivity for firms with access to better road infrastructure, highlighting the moderating effect of road infrastructure in ensuring gains from trade liberalization for firms.

Table 4: Baseline results

Dependent variable:	log TFP_{ijrt}				
	(1)	(2)	(3)	(4)	(5)
$\text{Input-tariff}_{ijrt}$	-0.006**		-0.005		-0.005
	(0.003)		(0.003)		(0.003)
$\text{Output-tariff}_{jt}$		-0.013		-0.008	-0.007
		(0.010)		(0.010)	(0.011)
$\text{Input-tariff}_{ijrt}\times\text{Infrastructure}_{rt}$			-0.006*		-0.005
			(0.003)		(0.003)
$\text{Output-tariff}_{jt}\times\text{Infrastructure}_{rt}$				-0.016***	-0.015**
				(0.006)	(0.006)
Observations	7463	7463	7463	7463	7463
Adjusted R^2	0.755	0.755	0.755	0.756	0.756
Firm FE	✓	✓	✓	✓	✓
Town-time FE	✓	✓	✓	✓	✓
Firm-time controls	✓	✓	✓	✓	✓

Notes: Data are for the years 1998 through 2009. Input and Output tariffs are constructed as outlined in Section 3.1.2. Variable Infrastructure refers to distance covered in 1hr travel time around firm node. Firm-time controls include exporter and FDI dummies and firm age. Standard errors in parenthesis are clustered at the sector-year level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 3: The effect of output tariffs moderated by the quality of infrastructure



Notes: This figure plots the estimated marginal effect of output tariffs on log TFP_{ijrt} from column (5) of Table 4 (on the vertical axis) as a function of $\text{Infrastructure}_{rt}$ (on the horizontal axis). Using the notation in equation (12), the point estimate plotted as a solid black line in the figure is given by $\hat{\beta} + \hat{\gamma} \times \text{Infrastructure}_{rt}$. Corresponding confidence intervals at the 90% level of statistical significance have been estimated for each value of $\text{Infrastructure}_{rt}$ in the estimation sample of 7463 observations used in Table 4. The figure focuses on values of $\text{Infrastructure}_{rt}$ from its 25^{th} percentile to its maximum value within the estimation sample. Relevant estimation sample values of the variable $\text{Infrastructure}_{rt}$ are indicated on the horizontal axis together with its mean computed on the entire sample at our disposal corresponding to the whole population of Ethiopian firms (pop mean). Notice that there are no observations in our estimation sample with values of $\text{Infrastructure}_{rt}$ between 7.06 and 7.95. As a consequence 66% of our sample observations have a value of $\text{Infrastructure}_{rt}$ which is above 7.95 and 33% of them have a value below 7.06. The marginal effect of output tariff computed at the 75^{th} percentile, which is not included in the diagram for the sake of space, is equal to -0.020.

4.2 Identification and Robustness

In this section, we address endogeneity concerns in our key independent variables using an instrumental variable estimation strategy. We then test for robustness of our results to alternate productivity and road infrastructure measures. Our results qualitatively support baseline results.

4.2.1 Instrumental Variables Estimation

As we argue in Section 3.3.1, the quality of road infrastructure might be endogeneous to firm productivity. To tackle this concern, we use instrument for road infrastructure in the interaction terms with tariffs using the slope of the terrain in the district that the firm is located in, weighted by distance to Addis, the capital city. Our IV is composed of geographical factors that we argue are largely exogeneous. We present results in Table 5. First-stage statistics reported in the table assure us that our instrument is strong and we find that it is negatively related to our road infrastructure variable as expected. From Columns (1) through (3), we find that second-stage results closely match our baseline results.

Table 5: Instrumenting for Road Infrastructure

Dependent variable:	log TFP$_{ijrt}$		
	(1)	(2)	(3)
Input-tariff$_{ijrt}$	-0.005*		-0.005*
	(0.003)		(0.003)
Output-tariff$_{jt}$		-0.007	-0.006
		(0.010)	(0.010)
Input-tariff$_{ijrt}$×Infrastructure$_{rt}$	-0.005		-0.003
	(0.003)		(0.003)
Output-tariff$_{jt}$×Infrastructure$_{rt}$		-0.017**	-0.016**
		(0.007)	(0.007)
Observations	7162	7162	7162
Adjusted R^2	0.753	0.754	0.754
Firm FE	√	√	√
Town-time FE	√	√	√
Firm-time controls	√	√	√
KP LM stat	59.824	68.638	56.729
P-val	0.000	0.000	0.000
KP F stat	1736.41	2269.115	866.624
SW F stat for input-t×infra			1744.401
SW F stat for output-t×infra			2209.529

Notes: Data are for the years 1998 through 2009. Input and Output tariffs are constructed as outlined in Section 3.1.2. Variable Infrastructure refers to distance covered in 1hr travel time around firm node. Firm-time controls include exporter and FDI dummies and firm age. The instrument for Infrastructure is terrain slope weighted by distance to Addis. Standard errors in parenthesis are clustered at the sector-year level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The Kleibergen-Paap (KP) LM statistic and related p-value allow to test for under-identification under the assumption of heteroskedasticity. Weak identification is tested with the KP Wald F test. In column (3) the KP Wald F statistic tests for weak identification of the two exogenous instruments jointly (Kleibergen and Paap, 2006). The Sanderson-Windmeijer (SW) tests allow to test for weak identification of each instrument separately (Sanderson and Windmeijer, 2016).

Next, we tackle the endogeneity of tariffs as discussed in Section 3.3.2. Though we find little evidence for tariffs being correlated with firm productivity or key industry characteristics, we check to see if the baseline relationship between tariffs and firm productivity remains robust to an instrumental variable estimation strategy. We hence estimate the direct effect of tariffs on firm productivity by instrumenting for long-run changes in tariffs with the pre-sample 1995 tariff level. Results are reported in Table 6. First-stage statistics reported in the table confirm the validity of our instrument. Second-stage results in Columns (1) through (3) show that though coefficients on the tariff variables are larger, they confirm the qualitative relationship between tariffs and firm productivity in our baseline results. Input tariff reductions are associated with increases in firm productivity, while output tariff reductions are not. Overall, our results are robust to accounting for endogeneity of key independent variables.

Table 6: Instrumenting for Tariffs

Dependent variable:	Δ log TFP$_{ijrt}$		
	(1)	(2)	(3)
Δ Input-tariff$_{ijrt}$	-0.891*		-1.106***
	(0.518)		(0.419)
Δ Output-tariff$_{jt}$		0.337	1.132
		(1.424)	(2.157)
Observations	1028	2177	1028
Town-time FE	✓	✓	✓
Firm-time controls	✓	✓	✓
KP LM stat	23.997	28.661	27.345
P-val	0.000	0.000	0.000
KP F stat	21.725	39.22	15.238
SW F stat for input-t			36.786
SW F stat for output-t			20.864

Notes: Data are for the years 1998 through 2009. Regressions are estimated in 5-year differences. Input and Output tariffs are constructed as outlined in Section 3.1.2. Instrument for tariffs is pre-sample tariff level in 1995. Firm-time controls include exporter and FDI dummies and firm age. Standard errors in parenthesis are clustered at the sector-year level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The Kleibergen-Paap (KP) LM statistic and related p-value allow to test for under-identification under the assumption of heteroskedasticity. Weak identification is tested with the KP Wald F test. In column (3) the KP Wald F statistic tests for weak identification of the two exogenous instruments jointly (Kleibergen and Paap, 2006). The Sanderson-Windmeijer (SW) tests allow to test for weak identification of each instrument separately (Sanderson and Windmeijer, 2016).

4.2.2 Robustness Checks

Finally, we undertake a battery of checks to ensure the robustness of our results to (1) alternative measures of the dependent variable; (2) alternative measures of road infrastructure; (3) alternative sub-samples.

First, we calculate alternative measures of firm productivity, to check if results are affected by the TFP estimation methods used and described in Section 3.2. In Column (1) of Table 7, we use a simple measure of productivity - labour productivity, measured as the share of real output to the number of employees. In columns (2) through (4), we estimate productivity using the Levinsohn-Petrin (LP) approach without the Ackerberg, Caves and Frazer (ACF) correction, accounting for a range of variables in the control function [13] and a one-step GMM method proposed by Wooldrigde (Wooldridge, 2009). Across the four Columns, we find that our results remain stable and consistent with the baseline.

Second, results may be affected by the choice of variable measuring road infrastructure. We use two alternate measures of road infrastructure. First, we use total area accessible by road in 1hr (computed using a buffer zone of 5 km on both sides of a road) as an alternative indicator of local accessibility and report results in Column (5). Second, in Column (6) we use a more comprehensive indicator that we draw from the O-D analysis, that measures improvements in travel time from each town to the town of Galafi (at the border with Djibouti, the main port for Ethiopian trade). With this variable, we capture an additional dimension of road connectivity,

[13] The control function approach is based on the methodology proposed by De Loecker et al. (2016), and consists of augmenting the set of variables affecting a firm's demand for materials. We do this by adding both input and output tariffs as well as the status of exporter of the firm.

particularly to the major port for imports. We are able to measure overall improvements in the road network and connectivity instead of our baseline focus on road improvements occurring in the immediate surroundings of a firm. Results in both Columns confirm our baseline results. Note that the coefficient on the interaction term between the output tariff and connectivity to Galafi is positive and statistically significant in Column (6) as expected, since greater distance to Galafi in this case measures lower connectivity.

Finally, we estimate our regressions without the year 2005, when the census was run as a representative survey, as discussed in Section 3.2. We also exclude years for which tariffs were imputed rather than directly obtained from the source. Our results in both cases remain qualitatively robust. Overall, our analysis in this section shows that our results are robust to alternate measures and samples, increasing confidence in our baseline results.

Table 7: Robustness to Alternative Measures and Samples

Productivity measure (in logs):	Y/L	No ACF	Cont Func	W-GMM	Baseline	
	(1)	(2)	(3)	(4)	(5)	(6)
Input-tariff$_{ijrt}$	-0.006	-0.004	-0.005	-0.003	-0.006*	-0.006*
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Output-tariff$_{ijrt}$	0.008	-0.009	-0.008	-0.004	-0.016	-0.021**
	(0.012)	(0.011)	(0.011)	(0.011)	(0.010)	(0.009)
Input-tariff$_{ijrt}$×Infrastructure$_{rt}$	0.000	-0.005	-0.004	-0.003		
	(0.003)	(0.003)	(0.003)	(0.003)		
Output-tariff$_{ijrt}$×Infrastructure$_{rt}$	-0.020**	-0.015**	-0.015**	-0.013*		
	(0.008)	(0.006)	(0.006)	(0.007)		
Input-tariff$_{ijrt}$×Infra-area$_{rt}$					-0.004	
					(0.007)	
Output-tariff$_{ijrt}$×Infra-area$_{rt}$					-0.062**	
					(0.030)	
Input-tariff$_{ijrt}$×Infra-Galafi$_{rt}$						-0.020
						(0.016)
Output-tariff$_{ijrt}$×Infra-Galafi$_{rt}$						0.107*
						(0.059)
Observations	7461	7463	7463	7463	7463	7463
Adjusted R^2	0.646	0.800	0.741	0.980	0.757	0.757
Firm FE	✓	✓	✓	✓	✓	✓
Town-time FE	✓	✓	✓	✓	✓	✓
Firm-time controls	✓	✓	✓	✓	✓	✓

Notes: Data are for the years 1998 through 2009. The dependent variables in Columns (1) through (4) are Labour Productivity and TFP measured by the LP method without ACF correction, a modified control function and the Wooldridge GMM approach respectively. Input and Output tariffs are constructed as outlined in Section 3.1.2. Variable Infrastructure refers to distance covered in 1hr travel time around firm node in Columns (1) through (4). Variable Infra-area refers to area covered in 1hr travel time around firm node. Variable Infra-Galafi refers to the travel distance to Galafi (border with Djibouti and port). Firm-time controls include exporter and FDI dummies and firm age. Standard errors in parenthesis are clustered at the sector-year level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5 Conclusion

In this paper, we examine the role of road infrastructure in moderating the effect of reductions in the input and output tariffs on the productivity of Ethiopian firms. We show that a reduction in the input tariff is associated with an increase in firm productivity. A reduction in the output tariff is associated with higher productivity for firms in towns with better road infrastructure. Our study confirms the existing finding in the literature that better access to intermediate inputs through a fall in the tariff can benefit domestic firms. In addition, it emphasizes a role for transport infrastructure in ensuring gains from trade. Lastly, in an environment where domestic intermediaries have market power, transmission of tariff reductions to domestic prices may not necessarily be stronger with better road infrastructure.

We belive that our analysis has implications for both trade and infrastructure policy in developing economies. While trade liberalization can improve firm performance by fostering competition and affording domestic firms better access to intermediate inputs, poor infrastructure can lead to weak transmission of tariff reductions to domestic prices, hampering gains from trade, particularly for remote regions, exacerbating concerns of regional inequality. Our study suggests that road infrastructure can complement the effects of trade liberalization on firm performance. Finally, greater competition in the intermediary sector can also yield better transmission of tariff reductions to domestic prices, ensuring that the benefits of trade are spread more uniformly.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Ahsan, Reshad N, "Input tariffs, speed of contract enforcement, and the productivity of firms in India," *Journal of International Economics*, 2013, *90* (1), 181–192.

Amiti, Mary and Amit K. Khandelwal, "Import Competition and Quality Upgrading," *Review of Economics and Statistics*, 2012, *95* (2), 476–490.

_ and Jozef Konings, "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia," *The American Economic Review*, 2007, *97* (5), 1611–1638.

Atkin, David and Dave Donaldson, "Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs," MIT, mimeo July 2015.

Bank, World, *Ethiopia Trade and Transformation - Diagniostic Trade Integration Study*, Washington, DC: World Bank, 2004.

Bas, Maria, "Input-trade liberalization and firm export decisions: Evidence from Argentina," *Journal of Development Economics*, 2012, *97* (2), 481 – 493.

Bigsten, Arne, Mulu Gebreeyesus, and Måns Söderbom, "Tariffs and Firm Performance in Ethiopia," *The Journal of Development Studies*, 2016, *52* (7), 986–1001.

Coşar, A. Kerem and Pablo D. Fajgelbaum, "Internal Geography, International Trade, and Regional Specialization," *American Economic Journal: Microeconomics*, February 2016, *8* (1), 24–56.

Coşar, Kerem and Banu Demir, "Domestic road infrastructure and international trade: Evidence from Turkey," *Journal of Development Economics*, 2016, *118*, 232 – 244.

De Loecker, Jan, "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica*, 2011, *79* (5), 1407–1451.

_ and Pinelopi K. Goldberg, "Firm Performance in a Global Market," *Annual Review of Economics*, 2014, *6* (1), 201–227.

_ , _ , Amit K. Khandelwal, and Nina Pavcnik, "Prices, Markups, and Trade Reform," *Econometrica*, 3 2016, *84* (2), 445–510.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *The American Economic Review*, forthcoming.

Duflo, Esther and Rohini Pande, "Dams," *The Quarterly Journal of Economics*, 2007, *122* (2), 601–646.

Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler, "The effects of structural reforms on productivity and profitability enhancing reallocation: evidence from Colombia," *Journal of Development Economics*, 2004, *75* (2), 333 – 371. 15th Inter American Seminar on Economics.

Goldberg, Pinelopi Koujianou, Amit Kumar Khandelwal, Nina Pavcnik, and Petia Topalova, "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *The Quarterly Journal of Economics*, 2010, *125* (4), 1727–1767.

Grossman, Gene M. and Elhanan Helpman, "Protection for Sale," *The American Economic Review*, 1994, *84* (4), 833–850.

Iimi, Atshushi, Haileyesus Adamtei, James Markland, and Eyasu Tsehaye, "Port Rail Connectivity and Agricultural Production Evidence from a Large Sample of Farmers in Ethiopia," World Bank Policy Research Working Paper 8088 2017.

Kleibergen, Frank and Richard Paap, "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, 2006, *133* (1), 97–126.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *The Review of Economic Studies*, 2003, *70* (2), 317–341.

Moller, Lars Christian, "Ethiopia's great run: the growth acceleration and how to pace it," Washington, D.C. : World Bank Group. 2015.

Mukim, Megha, "How Do Cities in Ethiopia Create Jobs?," World Bank Policy Research Working Paper 7866 2016.

Pavcnik, Nina, "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *The Review of Economic Studies*, 2002, *69* (1), 245–276.

Sanderson, Eleanor and Frank Windmeijer, "A Weak Instrument F-Test in Linear IV Models with Multiple Endogenous Variables," *Journal of Econometrics*, 2016, *190* (2), 212–221.

Shiferaw, Admasu, Måns Söderbom, Eyerusalem Siba, and Getnet Alemu, "Road Infrastructure and Enterprise Dynamics in Ethiopia," *The Journal of Development Studies*, 2015, *51* (11), 1541–1558.

Smeets, Valerie and Frederic Warzynski, "Estimating productivity with multi-product firms, pricing heterogeneity and the role of international trade," *Journal of International Economics*, 2013, *90* (2), 237 – 244.

Topalova, Petia and Amit K. Khandelwal, "Trade Liberalization and Firm Productivity: The Case of India," *The Review of Economics and Statistics*, 2011, *93* (3), 995–1009.

Volpe Martincus, Christian and Juan Blyde, "Shaky roads and trembling exports: Assessing the trade effects of domestic infrastructure using a natural experiment," *Journal of International Economics*, 2013, *90* (1), 148 – 161.

_ , Jerónimo Carballo, and Ana Cusolito, "Roads, exports and employment: Evidence from a developing country," *Journal of Development Economics*, 2017, *125*, 21 – 39.

Wang, Xu, Xiaobo Zhang, Zhuan Xie, and Yiping Huang, "Roads to Innovation: Firm-Level Evidence from China," IFPRI Discussion Paper 01542 2016.

Wooldridge, Jeffrey M, "On estimating firm-level production functions using proxy variables to control for unobservables," *Economics Letters*, 2009, *104* (3), 112–114.

Appendices

A Firm-level price index

Eslava et al. (2004) and Smeets and Warzynski (2013) propose an empirical model of the production function where firm-level revenues are used in the left hand side. Instead of deflating revenues with the standard vector of sector-level price indexes, these author propose a firm-level price index P_{it}. This Appendix discusses the procedure to adapt their methodology to the specificities of our data.

Step 1

First, we need to account for the fact that many products are not consistently identifiable across time due to the lack of a product category as identifier or to the fact that more than one product for the same firm in the same year have identical product code and unit measure. The solution we propose consists in treating non perfectly identifiable products as the elements of an aggregate product category that will be used alongside perfectly identifiable individual product categories. More precisely, all products with missing product code will be grouped in an aggregate product category (denoted with nim 'non identifiable missing') and all products with non-missing product code but still non identifiable (because they have the same values for both product code and unit measure within a firm-year pair) will be grouped in aggregate product categories depending on the non missing product code ($onih$ 'other non identifiable with product code h'). In order to aggregate the information contained in product-level observations we proceed as follows:

Step 1A

We derive a product-level price index as a weighted average of the prices of domestic sales and exports:

$$P_{hit} = \sum_{\nu=d,x} s_{hit}^{\nu} P_{hit}^{\nu} \tag{A-1}$$

where the superscripts d and x stand respectively for domestic and export market, and s_{hit}^{ν} is the share of the ν market in the total sales of product h by firm i at time t. We also perform some intuitive imputation in case some activity is reported but not all the required information is available. The approach we follow for inputing the data is in line with Eslava et al. (2004). In a nutshell: we compute sector-year level averages of P_{hit}^{ν} for $\nu \in \{d, x\}$ and we replace missing values of P_{hit}^{ν} with the respective sector-year average when we have a zero or missing value for sales or export quantity (value) and a non-missing, strictly positive value for sales or export value (quantity). Notice that when the value is missing (this is actually the minority of cases) the shares s_{hit}^{ν} cannot be computed. We correct for this by replacing the missing observation of value of domestic and/or export sales for a product-firm-time level observation with the average value of domestic and/or export sales across available observations of the same product in the same firm but in different years.

Step 1B

We compute $P_{nim,it}$ and $P_{onih,it}$ as the weighted average of P_{hit} for all h belonging to the respective group of non identifiable products, with weights computed as the h share of the total value (sales value plus export value) in the group. We create a database where products are actually product-aggregates but will be treated as individual products from now on.

Step 2

Second, we focus on product-level observations with perfectly identifiable products [8,174 product-level observations for 1984 firms]. We replicate Step 1A and we append the database created in Step1.

Then, we apply a Tornqvist formula to get the variation in firm-level prices. Notice that the dynamic structure of the Tornqvist formula requires that each product h is perfectly identifiable across time.

$$\Delta log(P_{it}) = \sum_{h} \frac{s_{hit} + s_{hi(t-1)}}{2} \times \left[log(P_{hit}) - log(P_{hi(t-1)})\right] \tag{A-2}$$

where s_{hit} is the share of product h total (both domestic and export) sales value over total sales value of the firm i at time t.

Finally, select as a base year the one where the number of active firms is at its maximum, i.e. the end of the time span (2009), we set $P_{i,2009} = 1$, and we proceed recursively (backward) to retrieve the firm-level prices:

$$log(P_{i(t-1)}) = log(P_{it}) - \Delta log(P_{it}) \ \forall t \leq 2009 \tag{A-3}$$

We first apply (A-3) only for those firm-year pairs (i, t) such that, for every year $t \leq k \leq 2009$, $\Delta log(P_{it})$ is non missing.

There are two potential computational caveats. First, a firm might not be observed in the base year. Consider the following example which illustrates the proposed solution. Take firm i and assume that the last year where it is observed is 2006. In that case the last $\Delta log(P_{it})$ that we can compute using (A-2) is $\Delta log(P_{i2006})$. We will set $log(P_{i2006})$ as the sector-level average for 2006, i.e. $\sum_j log(P_{j2006})/|J^{S(i)_{2006}}|$, where $|\cdot|$ is a cardinality operator, J^{k_t} is the set of firms j belonging to sector k for which we were able to retrieve $log(P_{jt})$, and $S(i)$ is the sector to which firm i belongs.

Second, a firm i might have a missing value for $\Delta log(P_{it})$ at a certain time t which is between two time intervals where it is potentially possible to apply the recursive formula (A-3). This would cause the breaking in the formula (this is the case for year 2004 and panel_id 6 for instance). Again, we solve this issue by replacing the missing observation of $log(P_{i(t-1)})$ with the sector-level average for that year.

We follow again the approach proposed by Eslava et al. (2004) to address these issues.